

# SEQUENCE LISTING

<110> Genencor International, Inc.  
Dunn-Coleman, N.  
Shetty, J.  
Duan, G.  
Sung, A.  
Qian, Y.

<120> A Method for the Preparation of a High Purity Rice Protein Concentrate

<130> GC830-PCT

<140> PCT/US05/05320

<141> 2005-02-17

<150> US 60/547,153

<151> 2004-02-23

<160> 11

<170> PatentIn version 3.2

<210> 1

<211> 2103

<212> DNA

<213> Humicola grisea var. thermoidea

<400> 1

atgcatacct	tctccaagct	cctcgctcctg	ggctctgccc	tccagtctgc	cctcggggcg	60
cctcacggct	cttcgcgtct	ccaggaacgc	gctgccgttg	ataccttcat	caacaccgag	120
aagcccacgc	catggaacaa	gctgctcgcc	aacatcggcc	ctaacggcaa	agccgctccc	180
gggtgccg	ccggcggtgt	gattgccagc	ccttccagga	cggaccctcc	ttgtacgtgg	240
tggcatggaa	tggacccaag	agactggttt	tagatgaaag	agagtttctg	ctaaccgcca	300
caccagact	tcttcacctg	gacccgcgat	gccgccctgg	tcctcaccgg	catcatcgag	360
tcccttgccc	acaactacaa	caccacctg	cagaccgtca	tccagaacta	cgtcgcgtcg	420
caggccaagc	tgcagcaggt	ctcgaacccc	tcgggaacct	tcgccgacgg	ctcgggtctc	480
gggtgaggcca	agttcaatgt	cgacctcact	gccttcactg	gcgaatgggg	tcgccctcag	540
agggacggcc	cgcccctg	cgccatcgct	ctcatccagt	acgccaagtg	gctgatcgcc	600
aacggctaca	agagcacggc	caagagcgct	gtctggcccc	tcgtcaagaa	cgatctcgcc	660
tacacggccc	agtactggaa	cgagaccggc	ttcgatctct	gggaggaggt	ccccggcagc	720
tcgttcttta	ccatcgccag	ctctcacagg	ggtagtcat	ttattgttca	gtgttttctc	780
attgaataat	taccggaatg	ccactgacgc	caaacagctc	tgactgaggg	tgcttacctc	840
gccgctcagc	tcgacaccga	gtgccgcgcc	tgacgacgc	tcgcccctca	ggttctgtgc	900
ttccagcagg	ccttctggaa	ctccaagggc	aactatgtcg	tctccaacag	taagatccct	960
acaccaacaa	aaaaaatcga	aaaggaacgt	tagctgacct	ttctagtcaa	cggcggcgag	1020
tatcgctccg	gcaaggacgc	caactcgatc	ctggcgctca	tccacaactt	cgaccctgag	1080
gccggctg	acaacctgac	cttccagccc	tgacgcgagc	gcgccctggc	caaccacaag	1140
gcctatgtcg	actcgttccg	caacctctac	gccatcaaca	agggcatcgc	ccagggcaag	1200
gccgttgccg	tcggccgcta	ctcggaggat	gtctactaca	acggcaaccc	gtggtacctg	1260
gccaactttg	ccggccgcca	gcagctctac	gacgccatct	acgtgtggaa	caagcagggc	1320
tccatcaccg	tgacctcggt	ctccctgccc	ttcttccgcg	accttgctctc	gtcggctcagc	1380
accggcacct	actccaagag	cagctcgacc	ttcaccaaca	tcgtcaacgc	cgtcaaggcc	1440
tacgccgacg	gcttcatcga	gggtggcgcc	aagtacaccc	cgtccaacgg	cgcgctcgcc	1500
gagcagtacg	accgcaacac	gggcaagccc	gactcggccg	ccgacctgac	gtggtcgtac	1560
tcggccttcc	tctcggccat	cgaccgcgcg	gcgggtctcg	tccccccgag	ctggcggggc	1620

```

agcgtggcca agagccagct gccgtccacc tgctcgcgca tgcagggtcgc cggcacctac 1680
gtcgccgcca cgagcacctc gttcccgtcc aagcagaccc cgaacccctc cgcggcgccc 1740
tccccgtccc cctacccgac cgccctgcgc gacgctagcg aggtgtacgt caccttcaac 1800
gagcgcgtgt cgaccgcgtg gggcgagacc atcaaggtgg tgggcaacgt gccggcgctg 1860
gggaactggg acacgtccaa ggcggtgacc ctgtcggcca gcgggtacaa gtcgaatgat 1920
cccctctgga gcatcacggg gcccatcaag gcgacgggct cggccgtgca gtacaagtat 1980
atcaaggtcg gcaccaacgg gaagattact tgggagtcgg accccaacag gagcattacc 2040
ctgcagacgg cgtcgtctgc gggcaagtgc gccgcgcaga cggatgaatga ttcgtggcgt 2100
taa 2103

```

```

<210> 2
<211> 634
<212> PRT
<213> Humicola grisea var. thermoidea

```

```

<400> 2
Met His Thr Phe Ser Lys Leu Leu Val Leu Gly Ser Ala Val Gln Ser
1 5 10 15
Ala Leu Gly Arg Pro His Gly Ser Ser Arg Leu Gln Glu Arg Ala Ala
20 25 30
Val Asp Thr Phe Ile Asn Thr Glu Lys Pro Ile Ala Trp Asn Lys Leu
35 40 45
Leu Ala Asn Ile Gly Pro Asn Gly Lys Ala Ala Pro Gly Ala Ala Ala
50 55 60
Gly Val Val Ile Ala Ser Pro Ser Arg Thr Asp Pro Pro Tyr Phe Phe
65 70 75 80
Thr Trp Thr Arg Asp Ala Ala Leu Val Leu Thr Gly Ile Ile Glu Ser
85 90 95
Leu Gly His Asn Tyr Asn Thr Thr Leu Gln Thr Val Ile Gln Asn Tyr
100 105 110
Val Ala Ser Gln Ala Lys Leu Gln Gln Val Ser Asn Pro Ser Gly Thr
115 120 125
Phe Ala Asp Gly Ser Gly Leu Gly Glu Ala Lys Phe Asn Val Asp Leu
130 135 140
Thr Ala Phe Thr Gly Glu Trp Gly Arg Pro Gln Arg Asp Gly Pro Pro
145 150 155 160
Leu Arg Ala Ile Ala Leu Ile Gln Tyr Ala Lys Trp Leu Ile Ala Asn
165 170 175
Gly Tyr Lys Ser Thr Ala Lys Ser Val Val Trp Pro Val Val Lys Asn
180 185 190
Asp Leu Ala Tyr Thr Ala Gln Tyr Trp Asn Glu Thr Gly Phe Asp Leu
195 200 205
Trp Glu Glu Val Pro Gly Ser Ser Phe Phe Thr Ile Ala Ser Ser His
210 215 220
Arg Ala Leu Thr Glu Gly Ala Tyr Leu Ala Ala Gln Leu Asp Thr Glu
225 230 235 240
Cys Arg Ala Cys Thr Thr Val Ala Pro Gln Val Leu Cys Phe Gln Gln
245 250 255
Ala Phe Trp Asn Ser Lys Gly Asn Tyr Val Val Ser Asn Ile Asn Gly
260 265 270
Gly Glu Tyr Arg Ser Gly Lys Asp Ala Asn Ser Ile Leu Ala Ser Ile
275 280 285
His Asn Phe Asp Pro Glu Ala Gly Cys Asp Asn Leu Thr Phe Gln Pro
290 295 300
Cys Ser Glu Arg Ala Leu Ala Asn His Lys Ala Tyr Val Asp Ser Phe
305 310 315 320
Arg Asn Leu Tyr Ala Ile Asn Lys Gly Ile Ala Gln Gly Lys Ala Val

```



				85					90					95			
Gly	Thr	Phe	Ala	Asp	Gly	Ser	Gly	Leu	Gly	Glu	Ala	Lys	Phe	Asn	Val		
			100					105					110				
Asp	Leu	Thr	Ala	Phe	Thr	Gly	Glu	Trp	Gly	Arg	Pro	Gln	Arg	Asp	Gly		
		115					120					125					
Pro	Pro	Leu	Arg	Ala	Ile	Ala	Leu	Ile	Gln	Tyr	Ala	Lys	Trp	Leu	Ile		
	130					135					140						
Ala	Asn	Gly	Tyr	Lys	Ser	Thr	Ala	Lys	Ser	Val	Val	Trp	Pro	Val	Val		
145					150					155					160		
Lys	Asn	Asp	Leu	Ala	Tyr	Thr	Ala	Gln	Tyr	Trp	Asn	Glu	Thr	Gly	Phe		
			165					170						175			
Asp	Leu	Trp	Glu	Glu	Val	Pro	Gly	Ser	Ser	Phe	Phe	Thr	Ile	Ala	Ser		
		180						185					190				
Ser	His	Arg	Ala	Leu	Thr	Glu	Gly	Ala	Tyr	Leu	Ala	Ala	Gln	Leu	Asp		
	195						200					205					
Thr	Glu	Cys	Arg	Ala	Cys	Thr	Thr	Val	Ala	Pro	Gln	Val	Leu	Cys	Phe		
	210					215					220						
Gln	Gln	Ala	Phe	Trp	Asn	Ser	Lys	Gly	Asn	Tyr	Val	Val	Ser	Asn	Ile		
225					230					235					240		
Asn	Gly	Gly	Glu	Tyr	Arg	Ser	Gly	Lys	Asp	Ala	Asn	Ser	Ile	Leu	Ala		
			245					250					255				
Ser	Ile	His	Asn	Phe	Asp	Pro	Glu	Ala	Gly	Cys	Asp	Asn	Leu	Thr	Phe		
		260						265				270					
Gln	Pro	Cys	Ser	Glu	Arg	Ala	Leu	Ala	Asn	His	Lys	Ala	Tyr	Val	Asp		
	275						280				285						
Ser	Phe	Arg	Asn	Leu	Tyr	Ala	Ile	Asn	Lys	Gly	Ile	Ala	Gln	Gly	Lys		
	290					295				300							
Ala	Val	Ala	Val	Gly	Arg	Tyr	Ser	Glu	Asp	Val	Tyr	Tyr	Asn	Gly	Asn		
305					310					315					320		
Pro	Trp	Tyr	Leu	Ala	Asn	Phe	Ala	Ala	Ala	Glu	Gln	Leu	Tyr	Asp	Ala		
			325					330						335			
Ile	Tyr	Val	Trp	Asn	Lys	Gln	Gly	Ser	Ile	Thr	Val	Thr	Ser	Val	Ser		
		340						345					350				
Leu	Pro	Phe	Phe	Arg	Asp	Leu	Val	Ser	Ser	Val	Ser	Thr	Gly	Thr	Tyr		
		355					360					365					
Ser	Lys	Ser	Ser	Ser	Thr	Phe	Thr	Asn	Ile	Val	Asn	Ala	Val	Lys	Ala		
	370					375				380							
Tyr	Ala	Asp	Gly	Phe	Ile	Glu	Val	Ala	Ala	Lys	Tyr	Thr	Pro	Ser	Asn		
385					390					395					400		
Gly	Ala	Leu	Ala	Glu	Gln	Tyr	Asp	Arg	Asn	Thr	Gly	Lys	Pro	Asp	Ser		
			405					410					415				
Ala	Ala	Asp	Leu	Thr	Trp	Ser	Tyr	Ser	Ala	Phe	Leu	Ser	Ala	Ile	Asp		
		420						425				430					
Arg	Arg	Ala	Gly	Leu	Val	Pro	Pro	Ser	Trp	Arg	Ala	Ser	Val	Ala	Lys		
		435					440					445					
Ser	Gln	Leu	Pro	Ser	Thr	Cys	Ser	Arg	Ile	Glu	Val	Ala	Gly	Thr	Tyr		
	450					455				460							
Val	Ala	Ala	Thr	Ser	Thr	Phe	Pro	Ser	Lys	Gln	Thr	Pro	Asn	Pro			
465					470				475					480			
Ser	Ala	Ala	Pro	Ser	Pro	Ser	Pro	Tyr	Pro	Thr	Ala	Cys	Ala	Asp	Ala		
			485					490					495				
Ser	Glu	Val	Tyr	Val	Thr	Phe	Asn	Glu	Arg	Val	Ser	Thr	Ala	Trp	Gly		
		500						505				510					
Glu	Thr	Ile	Lys	Val	Val	Gly	Asn	Val	Pro	Ala	Leu	Gly	Asn	Trp	Asp		
	515					520						525					
Thr	Ser	Lys	Ala	Val	Thr	Leu	Ser	Ala	Ser	Gly	Tyr	Lys	Ser	Asn	Asp		
	530					535					540						

Pro Leu Trp Ser Ile Thr Val Pro Ile Lys Ala Thr Gly Ser Ala Val  
545 550 555 560  
Gln Tyr Lys Tyr Ile Lys Val Gly Thr Asn Gly Lys Ile Thr Trp Glu  
565 570 575  
Ser Asp Pro Asn Arg Ser Ile Thr Leu Gln Thr Ala Ser Ser Ala Gly  
580 585 590  
Lys Cys Ala Ala Gln Thr Val Asn Asp Ser Trp Arg  
595 600

<210> 4  
<211> 10739  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> pTrex3g\_N13 plasmid

<400> 4  
aagcttacta gtactttctcg agctctgtac atgtccggtc gcgacgtacg cgtatcgatg 60  
gcgccagctg caggcggccg cctgcagcca cttgcagtcc cgtggaattc tcacggtgaa 120  
tgtaggcctt ttgtagggtg ggaattgtca ctcaagcacc cccaacctcc attacgcctc 180  
ccccatagag ttcccaatca gtgagtcacg gcaactgtct caaatagatt ggggagaagt 240  
tgacttccgc ccagagctga aggtcgcaca accgcatgat atagggtcgg caacggcaaa 300  
aaagcacgtg gctcaccgaa aagcaagatg tttgcgatct aacatccagg aacctggata 360  
catccatcat cacgcacgac cactttgatc tgctggtaaa ctctgattcg ccctaaaccg 420  
aagtgcgtgg taaatctaca cgtgggcccc tttcgggata ctgcgtgtgt cttctctagg 480  
tgccattctt ttcccttccct ctagtgttga attgtttgtg ttggagtccg agctgtaact 540  
acctctgaat ctctggagaa tgggtggacta acgactaccg tgcacctgca tcatgtatat 600  
aatagtgatc ctgagaaggg ggggttggag caatgtggga ctttgatggg catcaaacia 660  
agaacgaaga cgcctctttt gcaaagtttt gtttcgggta cgggtgaagaa ctggatactt 720  
gttgtgtctt ctgtgtattt ttgtggcaac aagaggccag agacaatcta ttcaaacacc 780  
aagcttgcctc ttttgagcta caagaacctg tgggggtatat atctagagtt gtgaagtcgg 840  
taatcccgct gtatagtaat acgagtcgca tctaaatact ccgaagctgc tgcgaaccgg 900  
gagaatcgag atgtgctgga aagcttctag cgagcggcta aattagcatg aaaggctatg 960  
agaaattctg gagacggctt gttgaatcat ggcgttccat tcttcgacaa gcaaagcggt 1020  
ccgtcgcagt agcaggcact cattcccga aaaactcgga gattcctaag tagcgatgga 1080  
accggaataa tataatagga aatacattga gttgcctcga cggttgcaat gcaggggtac 1140  
tgagcttgga cataactgtt ccgtacccca cctcttctca acctttggcg tttccctgat 1200  
tcagcgtacc cgtacaagtc gtaatcacta ttaaccaga ctgaccggac gtgttttgcc 1260  
cttcatttgg agaaataatg tcattgcatg gtgtaatttg cctgcttgac cgactggggc 1320  
tgttcgaagc ccgaatgtag gattgttatc cgaactctgc tcgtagaggc atgttgtgaa 1380  
tctgtgtcgg gcaggacacg cctcgaaggt tcacggcaag ggaaaccacc gatagcagtg 1440  
tctagtagca acctgtaaag ccgcaatgca gcatcactgg aaaatacaaa ccaatggcta 1500  
aaagtacata agttaatgcc taaagaagtc atataccagc ggctaataat tgtacaatca 1560  
agtggctaaa cgtaccgtaa tttgccaacg gcttgtgggg ttgcagaagc aacggcaaa 1620  
ccccacttcc ccacgtttgt ttcttcactc agtccaatct cagctgggtg tcccccaatt 1680  
gggtcgcttg tttgttccgg tgaagtgaag gaagacagag gtaagaatgt ctgactcgga 1740  
gcgttttgca tacaaccaag ggcagtgatg gaagacagtg aaatgttgac attcaaggag 1800  
tatttagcca gggatgcttg agtgatcgt gtaaggaggt catattgaaa tgtaaagtcg gcactgaaca 1860  
tactgtatag tcacttctga tgaagtgggtc catattgaaa tgtaaagtcg gcactgaaca 1920  
ggcaaaaagat tgagttgaaa ctgcctaaga tctcgggccc tcgggccttc ggcctttggg 1980  
tgtacatgtt tgtgctccgg gcaaatgcaa agtggtgtag gatcgaacac actgctgcct 2040  
ttaccaagca gctgagggtg tgtgataggc aaatgttcag gggccactgc atggtttcga 2100  
atagaaagag aagcttagcc aagaacaata gccgataaag atagcctcat taaacggaat 2160  
gagctagtag gcaaagtcag cgaatgtgta tatataaagg ttcgaggtcc gtgcctccct 2220  
catgctctcc ccactctact atcaactcag atcctccagg agacttgtag accatctttt 2280  
gaggcacaga aacccaatag tcaaccatca caagtttgta caaaaaagca ggctccgcgg 2340

ccgccccctt	caacatgcat	accttctcca	agctcctcgt	cctgggctct	gccgtccagt	2400
ctgccccctg	gcggcctcac	ggctcttcgc	gtctccagga	acgcgctgcc	gttgatacct	2460
tcatacaacac	cgagaagccc	atcgcatgga	acaagctgct	cgccaacatc	ggccctaacg	2520
gcaaagccgc	tcccgggtgcc	gccgccggcg	ttgtgattgc	cagcccttcc	aggacggacc	2580
ctccttgtag	gtgggtggcat	ggaatggacc	caagagactg	gttttagatg	aaagagagtt	2640
tctgctaacc	gccacaccca	gacttcttca	cctggaccgc	cgatgccgcc	ctggctctca	2700
ccggcatcat	cgagtccctt	ggccacaact	acaacaccac	cctgcagacc	gtcatccaga	2760
actacgtcgc	gtcgcaggcc	aagctgcagc	aggtctcgaa	ccctcggga	accttcgccg	2820
acggctcggg	tctcgggtgag	gccaagtcca	atgtcgacct	cactgccttc	actggcgaat	2880
ggggtcgccc	tcagagggac	ggcccggccc	tgcgcgccat	cgctctcatc	cagtagccca	2940
agtggctgat	cgccaacggc	tacaagagca	cggccaagag	cgctcgtctg	cccgtcgtca	3000
agaacgatct	cgcctacacg	gcccagtaact	ggaacgagac	cggcttcgat	ctctgggagg	3060
aggtcccccg	cagctcgttc	tttaccatcg	ccagctctca	caggggtgag	tcattttattg	3120
ttcagtggtt	tctcattgaa	taattaccgg	aatgccactg	acgccaaaca	gctctgactg	3180
agggtgctta	cctcgccgct	cagctcgaca	ccgagtgcgc	cgcttcgacg	accgtcgccc	3240
ctcaggttct	gtgcttccag	caggccttct	ggaactccaa	gggcaactat	gtcgtctcca	3300
acagtaagat	ccctacacca	acaaaaaaaa	tcgaaaagga	acgttagctg	accttcttag	3360
tcaacggcgc	cgagtatcgc	tccggcaagg	acgccaactc	gatcctggcg	tccatccaca	3420
acttcgaccc	tgaggccggc	tgcgacaacc	tgaccttcca	gccctgcagc	gagcgcgccc	3480
tggccaacca	caaggcctat	gtcgactcgt	tccgcaacct	ctacgccatc	aacaagggca	3540
tcgcccaggg	caaggccgtt	gccgtcggcc	gctactcgga	ggatgtctac	tacaacggca	3600
acccgtggta	cctggccaac	tttgccggcg	ccgagcagct	ctacgacgcc	atctacgtgt	3660
ggaacaagca	gggtcccatc	accgtgacct	cggctctccct	gcccttcttc	cgcgaccttg	3720
tctcgtcggg	cagcaccggc	acctaactcca	agagcagctc	gaccttcacc	aacatcgtca	3780
acgccgtcaa	ggcctacgcc	gacggcttca	tcgaggtggc	ggccaagtac	accccgcca	3840
acggcgcgct	cgccgagcag	tacgaccgca	acacgggcaa	gcccgaactc	gccgccgacc	3900
tgacgtggtc	gtactcggcc	ttcctctcgg	ccatcgaccg	ccgcgcgggt	ctcgtccccc	3960
cgagctggcg	ggccagcgtg	gccaagagcc	agctgcccgc	cacctgctcg	cgcacgcagg	4020
tcgcccggac	ctacgtcgcc	gccacgagca	cctcgttccc	gtccaagcag	accccgcaacc	4080
cctccgcggc	gccctccccg	tccccctacc	cgacggcctg	cgcggacgct	agcgaggtgt	4140
acgtcacctt	caacgagcgc	gtgtcgaccg	cgtggggcga	gaccatcaag	gtgggtgggca	4200
acgtgcccgc	gtgggggaac	tgggacacgt	ccaaggcggg	gacctgtcgc	gccagcgggt	4260
acaagtcgaa	tgatccccct	tggagcatca	cgggtgccat	caaggcgacg	ggctcggccg	4320
tgcagtacaa	gtatatcaag	gtcggcacca	acgggaagat	tacttgggag	tcggacccca	4380
acaggagcat	tacctgacg	acggcgctcg	ctgcccggca	gtgcgcgcgc	cagacgggtga	4440
atgattcgtg	gcgttaaaaag	gggtgggcgc	ccgaccagc	tttcttgtag	aaagtgggtga	4500
tcgcgccagc	tccgtgcgaa	agcctgacgc	accggtagat	tcttggtgag	cccgtatcat	4560
gacggcggcg	ggagctacat	ggccccgggt	gattttat	ttttgtatct	acttctgacc	4620
cttttcaaat	atacgggtcaa	ctcatcttcc	actggagatg	cggcctgctt	ggtattgcga	4680
tggtgtcagc	ttggcaaatt	gtggctttcg	aaaacacaaa	acgattcctt	agtagccatg	4740
cattttaaga	taacggaata	gaagaaagag	gaaattaaaa	aaaaaaaaaa	aacaaacatc	4800
ccgttcataa	cccgtagaat	cgcgcgtcct	cgtgtatccc	agtaccagtt	tattttgaat	4860
agctcgcccc	ctggagagca	tcctgaatgc	aagtaacaac	cgtagagggt	gacacggcag	4920
gtggtgctag	ggagcgtcgt	gttctacaag	gccagacgtc	ttcgcgggtg	atatatatgt	4980
atggttgact	gcaggctgct	cagcgacgac	agtcaagttc	gccctcgctg	cttgtgcaat	5040
aatcgagtg	gggaagccac	accgtgactc	ccatcttcca	gtaaagctct	gttgggtgtt	5100
atcagcaata	cacgtaattt	aaactcgtaa	gcatggggct	gatagcttaa	ttaccgttta	5160
ccagtgccat	ggttctgcag	ctttccttgg	cccgtaaaaa	tcggcgaagc	cagccaatca	5220
ccagctaggc	accagctaaa	ccctataatt	agtctcttat	caacaccatc	cgctcccccg	5280
ggatcaatga	ggagaatgag	ggggatgcgg	ggctaaagaa	gcctacataa	ccctcatgcc	5340
aactcccagt	ttactactcgt	cgagccaaca	tcctgactat	aagctaacac	agaatgcctc	5400
aatcctggga	agaactggcc	gctgataagc	gcgcccgcct	cgcaaaaacc	atccctgatg	5460
aatggaaagt	ccagacgctg	cctgcggaag	acagcgttat	tgatttccca	aagaaatcgg	5520
ggatcctttc	agaggccgaa	ctgaagatca	cagaggcctc	cgctgcagat	cttgtgtcca	5580
agctggcggc	cggagagttg	acctcgggtg	aagttacgct	agcattctgt	aaacgggcag	5640
caatcgccca	gcagttagta	gggtcccctc	tacctctcag	ggagatgtaa	caacgccacc	5700
ttatgggact	atcaagctga	cgctggcttc	tgtgcagaca	aactgcgccc	acgagttctt	5760

ccctgacgcc	gctctcgcgc	aggcaagggg	actcgatgaa	tactacgcaa	agcacaagag	5820
acccggttgg	ccactccatg	gcctcccat	ctctctcaaa	gaccagcttc	gagtcaaggt	5880
acaccgttgc	ccctaagtcg	ttagatgtcc	ctttttgtca	gctaacatat	gccaccagg	5940
ctacgaaaca	tcaatgggct	acatctcatg	gctaaacaag	tacgacgaag	gggactcgg	6000
tctgacaacc	atgctccgca	aagccgggtg	cgtcttctac	gtcaagacct	ctgtcccgc	6060
gaccctgatg	gtctgcgaga	cagtcaacaa	catcatcggg	cgcaccgtca	acccacgcaa	6120
caagaactgg	tcgtgcggcg	gcagttctgg	tggtgaggg	gcgatcgttg	ggattcgcrv	6180
tggtggcgtc	atcgggttag	gaacggatat	cggtggtcgc	attcgagtgc	cggccgcgtt	6240
caacttcctg	tacggtctaa	ggccgagtca	tgggcggctg	ccgtatgcaa	agatggcgaa	6300
cagcatggag	ggtcaggaga	cggtcacacg	cgttgctggg	ccgattacgc	actctgttga	6360
gggtgagtc	ttcgctctt	cttctctttt	ctgctctata	ccaggcctcc	actgtcctcc	6420
tttcttgcct	tttatactat	atacgagacc	ggcagtcact	gatgaagtat	gttagacctc	6480
cgcctcttca	ccaaatccgt	cctcggtcag	gagccatgga	aatacgactc	caaggctatc	6540
cccatgccct	ggcgccagtc	cgagtcggac	attattgcct	ccaagatcaa	gaacggcggg	6600
ctcaatatcg	gctactacaa	cttcgacggc	aatgtccttc	cacaccctcc	tatcctgcgc	6660
ggcgtggaaa	ccaccgtcgc	cgcactcgcc	aaagccggtc	acaccgtgac	cccgtggacg	6720
ccatacaagc	acgatttcgg	ccacgatctc	atctcccata	tctacgcggc	tgacggcagc	6780
crvgccgacg	taatgcgcga	tatcagtgca	tccggcgagc	cggcgattcc	aaatatcaaa	6840
gacctactga	acccgaacat	caaagctggt	aacatgaacg	agctctggga	cacgcattctc	6900
cagaagtggg	attaccagat	ggagtacctt	gagaaatggc	gggaggctga	agaaaaggcc	6960
gggaagggaac	tggacgccat	catcgcgccg	attacgccta	ccgctgcggg	acggcatgac	7020
cagttccggt	actatgggta	tgctctctgt	atcaacctgc	tggatttcac	gagcgtgggt	7080
gttccgggtta	cctttgcgga	taagaacatc	gataagaaga	atgagagttt	caaggcgggt	7140
agtgagcttg	atgccctcgt	gcaggaagag	tatgatccgg	aggcgtacca	tggggcaccg	7200
gttgcaagtgc	aggttatcgg	acgggactc	agtgaagaga	ggacgttggc	gattgcagag	7260
gaagtgggga	agttgctggg	aaatgtgggt	actccatagc	taataagtgt	cagatagcaa	7320
tttgcaacaag	aatcaatac	cagcaactgt	aaataagcgc	tgaagtgacc	atgccatgct	7380
acgaaaagagc	agaaaaaac	ctgccgtaga	accgaagaga	tatgacacgc	ttccatctct	7440
caaaggaaga	atcccttcag	ggttgcggtt	ccagcttaga	cacgtataac	ggcacaagtg	7500
tctctcacca	aatgggttat	atctcaaatg	tgatctaagg	atggaaagcc	cagaatatcg	7560
atcgcgcgca	gatccatata	tagggcccgg	gttataatta	cctcaggtcg	acgtcccatg	7620
gccatttcgaa	ttcgtaatca	tggtcatagc	tgtttcctgt	gtgaaattgt	tatccgctca	7680
caatttcacac	caacatacga	gccggaagca	taaagtgtaa	agcctggggg	gcctaatagag	7740
tgagctaact	cacattaatt	gcgttgcgct	cactgcccgc	tttccagtcg	ggaaacctgt	7800
cgtgccagct	gcattaatga	atcggccaac	gcgcggggag	aggcggtttg	cgtattgggc	7860
gctcttccgc	ttctctgcct	actgactcgc	tgcgctcgg	cgttcggctg	cggcgagcgg	7920
tatcagctca	ctcaaaggcg	gtaatacgg	tatccacaga	atcaggggat	aacgcaggaa	7980
agaacatgtg	agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	8040
cgtttttcca	taggctccgc	ccccctgacg	agcatcacia	aaatcgacgc	tcaagtcaga	8100
ggtggcgaaa	cccagacagga	ctataaagat	accaggcggt	tccccctgga	agctccctcg	8160
tgcgctctcc	tgttccgacc	ctgccgctta	ccggatacct	gtccgccttt	ctcccttcgg	8220
gaagcgtggc	gctttctcat	agctcacgct	gtaggtatct	cagttcgggtg	taggtcggtc	8280
gctccaagct	gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	8340
gtaactatcg	tcttgagtcc	aaccgggtaa	gacacgactt	atcgccactg	gcagcagcca	8400
ctggtaacag	gattagcaga	gcgaggtatg	taggcgggtg	tacagagttc	ttgaagtgg	8460
ggcctaacta	cggctacact	agaagaacag	tatttggtat	ctgcgctctg	ctgaagccag	8520
ttaccttcgg	aaaaagagtt	ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	8580
gtggtttttt	tgtttgcaag	cagcagatta	cgcgcagaaa	aaaaggatct	caagaagatc	8640
ctttgatctt	ttctacgggg	tctgacgctc	agtgaacga	aaactcacgt	taagggaatt	8700
tggtcatgag	attatcaaaa	aggatcttca	cttagatcct	tttaaattaa	aaatgaagtt	8760
ttaaatcaat	ctaaagtata	tatgagtaaa	cttggtctga	cagttaccaa	tgcttaatca	8820
gtgaggcacc	tatctcagcg	atctgtctat	ttcgttcata	catagttgcc	tgactccccg	8880
tcgtgtagat	aactacgata	cgggagggct	taccatctgg	ccccagtgt	gcaatgatac	8940
cgcgagaccc	acgctcaccg	gctccagatt	tatcagcaat	aaaccagcca	gccggaaggg	9000
ccgagcgcg	aagtggctct	gcaactttat	ccgcctccat	ccagtctatt	aattgttgcc	9060
gggaagctag	agtaagtagt	tcgccagtta	atagtttgcg	caacgttggt	gccattgcta	9120
caggcatcgt	ggtgtcacgc	tcgtcggttg	gtatggcttc	attcagctcc	ggttcccaac	9180

gatcaaggcg	agttacatga	tcccccatgt	tgtgcaaaaa	agcgggttagc	tccttcggtc	9240
ctccgatcgt	tgtcagaagt	aagttggccg	cagtgttatc	actcatggtt	atggcagcac	9300
tgcataattc	tcttactgtc	atgccatccg	taagatgctt	ttctgtgact	ggtagagtact	9360
caaccaagtc	attctgagaa	tagtgtatgc	ggcgaccgag	ttgctcttgc	ccggcgtaa	9420
tacgggataa	taccgcgcca	catagcagaa	ctttaaaagt	gctcatcatt	ggaaaacgtt	9480
cttcggggcg	aaaactctca	aggatcttac	cgctgttgag	atccagttcg	atgtaaccca	9540
ctcgtgcacc	caactgatct	tcagcatctt	ttactttcac	cagcgtttct	gggtgagcaa	9600
aaacaggaag	gcaaaatgcc	gcaaaaaagg	gaataagggc	gacacggaaa	tgttgaatac	9660
tcatactctt	cctttttcaa	tattattgaa	gcatttatca	gggttattgt	ctcatgagcg	9720
gatacatatt	tgaatgtatt	tagaaaaata	aacaaatagg	ggttccgcgc	acattttccc	9780
gaaaagtgcc	acctgacgtc	taagaaacca	ttattatcat	gacattaacc	tataaaaaata	9840
ggcgtatcac	gaggcccttt	cgtctcgcgc	gtttcgggtga	tgacgggtgaa	aacctctgac	9900
acatgcagct	cccggagacg	gtcacagctt	gtctgtaagc	ggatgccggg	agcagacaag	9960
cccgtcaggg	cgctcagcg	gggtgtggcg	gggtgcgggg	ctggcttaac	tatgcggcat	10020
cagagcagat	tgtactgaga	gtgcaccata	aaattgtaaa	cgtaaatatt	ttgttaaaat	10080
tcgcgttaaa	tttttgttaa	atcagctcat	tttttaacca	ataggccgaa	atcggcaaaa	10140
tcccttataa	atcaaaagaa	tagccccgaga	taggggtgag	tgttggtcca	gtttggaaca	10200
agagtccact	attaaagaac	gtggactcca	acgtcaaagg	gcgaaaaacc	gtctatcagg	10260
gcgatggccc	actacgtgaa	ccatcaccca	aatcaagttt	tttggggctg	aggtgccgta	10320
aagcactaaa	tcggaaccct	aaagggagcc	cccgathtag	agcttgacgg	ggaaagccgg	10380
cgaacgtggc	gagaaaggaa	gggaagaaag	cgaagggagc	gggcgctagg	gcgctggcaa	10440
gtgtagcgg	cacgctgcgc	gtaaccacca	caccgcgcgc	gcttaatgcg	ccgctacagg	10500
gcgcgtacta	tggttgcttt	gacgtatgcg	gtgtgaaata	ccgcacagat	gcgtaaggag	10560
aaaataccgc	atcaggcgcc	attcgccatt	caggctgcgc	aactgttggg	aagggcgatc	10620
gggtgcgggc	tcttcgctat	tacgccagct	ggcgaaagg	ggatgtgctg	caaggcgatt	10680
aagttgggta	acgccaggg	tttcccagtc	acgacgttgt	aaaacgacgg	ccagtgcgc	10739

<210> 5

<211> 2169

<212> DNA

<213> *Aspergillus awamori* var. *kawachi*

<400> 5

atgtcgttcc	gatctcttct	cgccctgagc	ggccttgtct	gctcgggggt	ggcaagtgtg	60
atttccaagc	gcgcgacctt	ggattcgtgg	ttgagcaacg	aagcgaccgt	ggcccgtact	120
gcgatcctga	ataacatcgg	ggcggacgg	gcttgggtgt	cgggcgcgga	ctctggcatt	180
gtcgttgcca	gtcccagcac	cgataaccgc	gactgtatgt	tttgagttcg	gattatgaat	240
gtgtcttgg	tgattgatgc	tgactggcgt	gtcttttgat	gattgtagac	ttctacacct	300
ggactcgcga	ctctgggtct	gtcatcaaga	ccctcgtcga	cctcttcgc	aatggagata	360
ctgatctcct	ttccaccatt	gagcactaca	tctcctctca	ggcaattatt	caggggtgtca	420
gtaacccctc	tgggtgatctg	tccagcgggtg	gtcttggtga	gccaagtctc	aatgtcgatg	480
agactgccta	caccgggttct	tggggacggc	cgcagcgtga	tggctctgcc	ctgagagcaa	540
ctgctatgat	cggctttggg	cagtggctgc	ttgtatgttc	tccacctcct	tgcgtctgat	600
ctgcaacata	tgtagccgac	tggtcaggac	aatggctaca	ccagcgtgc	aacagagatt	660
gttttgcccc	tcgttaggaa	cgacctgtcg	tatgtggctc	agtactggaa	ccagacggga	720
tatgggtgtg	ttgattgatc	ggggttcaag	gggttttgtg	catcggagct	aacttcgcgg	780
tcgcagatct	ctgggaagaa	gttaatggct	cgtccttctt	cactattgcc	gtgcaacacc	840
gcgccctcgt	cgaaggtagt	gccttcgcga	cggccgtcgg	ctcgtcctgc	tcctgggtgtg	900
attcgcaggc	acctcagatt	ctctgttact	tgagtcctt	ctggaccggc	agctacatcc	960
tggccaactt	tgacagcagc	cgttcgggca	aggacacaaa	cacctcctg	ggaagcatcc	1020
acacctttga	tcctgaggct	ggatgcgacg	actccacctt	ccagccctgc	tccccgcgtg	1080
cgctcgccaa	ccataaggag	gttgtagact	ctttccgctc	gatctatact	ctcaacgatg	1140
gtctcagtga	cagtgaaggc	gttgcggtcg	gtcgggtacc	tgaggatagc	tactacaacg	1200
gcaaccctgt	gttcctgtgc	accttggctg	ccgcggaaca	gctgtacgat	gctctgtacc	1260
agtgggacaa	gcaggggtcg	ttggagatca	cagacgtgtc	acttgacttc	ttcaaggctc	1320
tgtacagtgg	tgtgccacc	ggcacgtact	cttcgtccag	ctcgacctat	agcagcattg	1380
tgagtgccgt	caagactttc	gctgatgggt	ttgtttctat	tgtggttaagt	ctacgctaga	1440



cgagcgcctca	tatttacaga	gggtgcgtac	taacaggatt	aggaaactca	cgccgcaagc	1500
aacggctctc	tgtctgagca	attcgacaag	tctgatggcg	acgagctttc	tgctcgcgat	1560
ctgacctggg	cttacgctgc	tctgctgacc	gccaacaacc	gtcgtaatte	tgctgtgccc	1620
ccgtcttggg	gtgagacctc	tgccagcagc	gtgcccggca	cctgtgcggc	tacctctgcc	1680
tctgggtacct	acagcagtg	gaccgtcacc	tcgtggccga	gcacgtgggc	tactggtggc	1740
accactacga	cggctactac	cactggatcg	ggcggcgtga	cctcgaccag	caagaccacc	1800
acaactgcta	gtaagaccag	caccactacg	tcctcgacct	cctgcaccac	ccccactgcc	1860
gtagctgtga	cctttgatct	gacggcgacc	accacctacg	gcgagaacat	ctacctggtc	1920
gggtcgatct	ctcagctcgg	tgactgggag	accagcgatg	gcatagctct	gagcgctgac	1980
aagtacactt	ccagcaaccc	gctttggtat	gtaactgtga	ctctgccggc	tggtgagtc	2040
tttgagtaca	agttcatccg	cgtcgagagc	gatgactccg	tggagtggga	gagcgaccgc	2100
aaccgggaat	acaccgttcc	tcaggcgctgc	ggcgagtcga	ccgcgacggt	gaccgacacc	2160
tggcggtag						2169

<210> 6  
 <211> 24  
 <212> PRT  
 <213> *Aspergillus awamori* var. *kawachi*

<400> 6  
 Met Ser Phe Arg Ser Leu Leu Ala Leu Ser Gly Leu Val Cys Ser Gly  
 1 5 10 15  
 Leu Ala Ser Val Ile Ser Lys Arg  
 20

<210> 7  
 <211> 615  
 <212> PRT  
 <213> *Aspergillus awamori* var. *kawachi*

<400> 7  
 Ala Thr Leu Asp Ser Trp Leu Ser Asn Glu Ala Thr Val Ala Arg Thr  
 1 5 10 15  
 Ala Ile Leu Asn Asn Ile Gly Ala Asp Gly Ala Trp Val Ser Gly Ala  
 20 25 30  
 Asp Ser Gly Ile Val Val Ala Ser Pro Ser Thr Asp Asn Pro Asp Tyr  
 35 40 45  
 Phe Tyr Thr Trp Thr Arg Asp Ser Gly Leu Val Ile Lys Thr Leu Val  
 50 55 60  
 Asp Leu Phe Arg Asn Gly Asp Thr Asp Leu Leu Ser Thr Ile Glu His  
 65 70 75 80  
 Tyr Ile Ser Ser Gln Ala Ile Ile Gln Gly Val Ser Asn Pro Ser Gly  
 85 90 95  
 Asp Leu Ser Ser Gly Gly Leu Gly Glu Pro Lys Phe Asn Val Asp Glu  
 100 105 110  
 Thr Ala Tyr Thr Gly Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala  
 115 120 125  
 Leu Arg Ala Thr Ala Met Ile Gly Phe Gly Gln Trp Leu Leu Asp Asn  
 130 135 140  
 Gly Tyr Thr Ser Ala Ala Thr Glu Ile Val Trp Pro Leu Val Arg Asn  
 145 150 155 160  
 Asp Leu Ser Tyr Val Ala Gln Tyr Trp Asn Gln Thr Gly Tyr Asp Leu  
 165 170 175  
 Trp Glu Glu Val Asn Gly Ser Ser Phe Phe Thr Ile Ala Val Gln His  
 180 185 190  
 Arg Ala Leu Val Glu Gly Ser Ala Phe Ala Thr Ala Val Gly Ser Ser  
 195 200 205



<220>  
 <223> primer  
  
 <400> 8  
 caacatgcat accttctcca agtcctc 28  
  
 <210> 9  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> primer  
  
 <400> 9  
 ttaacgccac gaatcattca ccgtc 25  
  
 <210> 10  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> primer  
  
 <400> 10  
 caccatgtcg ttccgatctc ttctc 25  
  
 <210> 11  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> primer  
  
 <400> 11  
 ctaccgccag gtgtcggta c 21